

SEQUENCE LISTING

<110> Genencor International, Inc.

Dunn-Coleman, N.

Shetty, J.

Duan, G.

Sung, A.

Qian, Y.

<120> A Method for the Preparation of a High Purity Rice Protein Concentrate

<130> GC830-PCT

<140> PCT/US05/05320

<141> 2005-02-17

<150> US 60/547,153

<151> 2004-02-23

<160> 11

<170> PatentIn version 3.2

<210> 1

<211> 2103

<212> DNA

<213> Humicola grisea var. thermoidea

<400> 1

atgcataacct	tctccaagct	cctcgccctg	ggctctgccg	tccagtctgc	cctcgcccg	60
cctcacggct	tttcgcgtct	ccaggaacgc	gctgccgttg	ataccttcat	caacaccgag	120
aagcccatcg	catggAACAA	gctgctcgcc	aacatcgccc	ctaacggcaa	agccgctccc	180
ggtgccgccc	ccggcgttgt	gattgccagc	ccttccagga	cggaccctcc	ttgtacgtgg	240
tggcatggaa	tggacccaag	agactggttt	tagatgaaag	agagtttctg	ctaaccgcca	300
caccagact	tcttcacctg	gaccgcgat	gccgcctgg	tcctcaccgg	catcatcgag	360
tcccttgccc	acaactacaa	caccaccctg	cagaccgtca	tccagaacta	cgtcgctcg	420
caggccaagc	tgcagcaggt	ctcgaacccc	tcggAACCT	tcgcccacgg	ctcggtctc	480
ggtgaggcca	agttcaatgt	cgacactact	gccttcaactg	gcgaatgggg	tcgcctcag	540
agggacggcc	cgcgcctg	cgccatcgct	ctcatccagt	acgccaagtg	gctgatcgcc	600
aacggctaca	agagcacggc	caagagcgtc	gtctggccc	tcgtcaagaa	cgatctcgcc	660
tacacggccc	agtaactggaa	cgagaccggc	ttcgatctct	gggaggaggt	ccccggcagc	720
tcgttctta	ccatcgccag	ctctcacagg	ggtgagtcat	ttattgttca	gtgtttctc	780
attgaataat	taccggaatg	ccactgacgc	caaacagctc	tgactgaggg	tgcttacctc	840
gccgctcagc	tcgacaccga	gtgccgcgc	tgcacgaccg	tcgcccctca	ggttctgtgc	900
ttccagcagg	ccttctggaa	ctccaagggc	aactatgtcg	tctccaacag	taagatccct	960
acaccaacaa	aaaaaatcga	aaaggaacgt	tagctgaccc	ttctagtcaa	cgccggcgag	1020
tatcgctccg	gcaaggacgc	caactcgatc	ctggcgatcca	tccacaactt	cgaccctgag	1080
gccggctcgc	acaacctgac	cttccagccc	tgcagcgagc	gcccctggc	caaccacaag	1140
gcctatgtcg	actcggtccg	caacctctac	gccatcaaca	agggcatcg	ccagggcaag	1200
gccgttgcgc	tcggccgcta	ctcgaggat	gtctactaca	acggcaaccc	gtggtacctg	1260
gccaactttg	ccggccgcga	gcagctctac	gacgccatct	acgtgtggaa	caagcagggc	1320
tccatcaccg	tgacctcggt	ctccctgccc	ttcttccgc	accttgcctc	gtcggtcagc	1380
accggcacct	actccaagag	cagctcgacc	ttcaccaaca	tcgtcaacgc	cgtcaaggcc	1440
tacgcccacg	gcttcatcga	ggtggcgccc	aagtacaccc	cgtccaacgg	cgcgctcgcc	1500
gagcagtacg	accgcaacac	gggcaagccc	gactcgccg	ccgacctgac	gtggtcgtac	1560
tcggcctcc	tctcgccat	cgaccgcgc	gccccggag	ctggcggcc		1620

agcgtggcca	agagccagct	gccgtccacc	tgctcgca	tcgaggctgc	cggcacctac	1680
gtcgccgcca	cgagcacctc	gttcccgtcc	aagcagaccc	cgaacccctc	cgcggcgccc	1740
tccccgtccc	cctaccggac	cgcctgcgcg	gacgctagcg	aggtgtacgt	caccttcaac	1800
gagcgcgtgt	cgaccgcgtg	gggcgagacc	atcaaggtgg	tgggcaacgt	gccggcgctg	1860
gggaactggg	acacgtccaa	ggcggtgacc	ctgtcgccca	gcgggtacaa	gtcaaatgat	1920
cccctctgga	gcatcacggt	gcccatcaag	gcgacgggct	cggccgtgca	gtacaagtat	1980
atcaaggtcg	gcaccaaacgg	gaagattact	tgggagtcgg	accccaacag	gagcattacc	2040
ctgcagacgg	cgtcgtctgc	ggcaagtgc	gccgcgcaga	cggtgaatga	ttcgtggcgt	2100
taa						2103

<210> 2
<211> 634
<212> PRT
<213> Humicola grisea var. thermoidea

<400> 2						
Met His Thr Phe Ser Lys Leu Leu Val	Leu Gly Ser Ala Val Gln Ser					
1 5	10 15					
Ala Leu Gly Arg Pro His Gly Ser Ser	Arg Leu Gln Glu Arg Ala Ala					
20 25	30					
Val Asp Thr Phe Ile Asn Thr Glu Lys	Pro Ile Ala Trp Asn Lys Leu					
35 40	45					
Leu Ala Asn Ile Gly Pro Asn Gly Lys	Ala Ala Pro Gly Ala Ala Ala					
50 55	60					
Gly Val Val Ile Ala Ser Pro Ser Arg	Thr Asp Pro Pro Tyr Phe Phe					
65 70	75 80					
Thr Trp Thr Arg Asp Ala Ala Leu Val	Leu Thr Gly Ile Ile Glu Ser					
85 90	95					
Leu Gly His Asn Tyr Asn Thr Thr Leu	Gln Thr Val Ile Gln Asn Tyr					
100 105	110					
Val Ala Ser Gln Ala Lys Leu Gln Gln	Val Ser Asn Pro Ser Gly Thr					
115 120	125					
Phe Ala Asp Gly Ser Gly Leu Gly Glu	Ala Lys Phe Asn Val Asp Leu					
130 135	140					
Thr Ala Phe Thr Gly Glu Trp Gly Arg	Pro Gln Arg Asp Gly Pro Pro					
145 150	155 160					
Leu Arg Ala Ile Ala Leu Ile Gln Tyr	Ala Lys Trp Leu Ile Ala Asn					
165 170	175					
Gly Tyr Lys Ser Thr Ala Lys Ser Val	Val Trp Pro Val Val Lys Asn					
180 185	190					
Asp Leu Ala Tyr Thr Ala Gln Tyr Trp	Asn Glu Thr Gly Phe Asp Leu					
195 200	205					
Trp Glu Glu Val Pro Gly Ser Ser	Phe Thr Ile Ala Ser Ser His					
210 215	220					
Arg Ala Leu Thr Glu Gly Ala Tyr Leu	Ala Ala Gln Leu Asp Thr Glu					
225 230	235 240					
Cys Arg Ala Cys Thr Thr Val Ala Pro	Gln Val Leu Cys Phe Gln Gln					
245 250	255					
Ala Phe Trp Asn Ser Lys Gly Asn	Tyr Val Val Ser Asn Ile Asn Gly					
260 265	270					
Gly Glu Tyr Arg Ser Gly Lys Asp	Ala Asn Ser Ile Leu Ala Ser Ile					
275 280	285					
His Asn Phe Asp Pro Glu Ala Gly Cys	Asp Asn Leu Thr Phe Gln Pro					
290 295	300					
Cys Ser Glu Arg Ala Leu Ala Asn His	Lys Ala Tyr Val Asp Ser Phe					
305 310	315 320					
Arg Asn Leu Tyr Ala Ile Asn Lys Gly	Ile Ala Gln Gly Lys Ala Val					

325	330	335
Ala Val Gly Arg Tyr Ser Glu Asp Val	Tyr Tyr Asn Gly Asn Pro Trp	
340	345	350
Tyr Leu Ala Asn Phe Ala Ala Ala	Glu Gln Leu Tyr Asp Ala Ile Tyr	
355	360	365
Val Trp Asn Lys Gln Gly Ser Ile Thr Val Thr Ser Val Ser Leu Pro		
370	375	380
Phe Phe Arg Asp Leu Val Ser Ser Val Ser Thr Gly Thr Tyr Ser Lys		
385	390	395
Ser Ser Ser Thr Phe Thr Asn Ile Val Asn Ala Val Lys Ala Tyr Ala		
405	410	415
Asp Gly Phe Ile Glu Val Ala Ala Lys Tyr Thr Pro Ser Asn Gly Ala		
420	425	430
Leu Ala Glu Gln Tyr Asp Arg Asn Thr Gly Lys Pro Asp Ser Ala Ala		
435	440	445
Asp Leu Thr Trp Ser Tyr Ser Ala Phe Leu Ser Ala Ile Asp Arg Arg		
450	455	460
Ala Gly Leu Val Pro Pro Ser Trp Arg Ala Ser Val Ala Lys Ser Gln		
465	470	475
Leu Pro Ser Thr Cys Ser Arg Ile Glu Val Ala Gly Thr Tyr Val Ala		
485	490	495
Ala Thr Ser Thr Ser Phe Pro Ser Lys Gln Thr Pro Asn Pro Ser Ala		
500	505	510
Ala Pro Ser Pro Ser Pro Tyr Pro Thr Ala Cys Ala Asp Ala Ser Glu		
515	520	525
Val Tyr Val Thr Phe Asn Glu Arg Val Ser Thr Ala Trp Gly Glu Thr		
530	535	540
Ile Lys Val Val Gly Asn Val Pro Ala Leu Gly Asn Trp Asp Thr Ser		
545	550	555
Lys Ala Val Thr Leu Ser Ala Ser Gly Tyr Lys Ser Asn Asp Pro Leu		
565	570	575
Trp Ser Ile Thr Val Pro Ile Lys Ala Thr Gly Ser Ala Val Gln Tyr		
580	585	590
Lys Tyr Ile Lys Val Gly Thr Asn Gly Lys Ile Thr Trp Glu Ser Asp		
595	600	605
Pro Asn Arg Ser Ile Thr Leu Gln Thr Ala Ser Ser Ala Gly Lys Cys		
610	615	620
Ala Ala Gln Thr Val Asn Asp Ser Trp Arg		
625	630	

<210> 3
 <211> 604
 <212> PRT
 <213> Humicola grisea var. thermoidea

<400> 3			
Ala Ala Val Asp Thr Phe Ile Asn Thr Glu Lys Pro Ile Ala Trp Asn			
1	5	10	15
Lys Leu Leu Ala Asn Ile Gly Pro Asn Gly Lys Ala Ala Pro Gly Ala			
20	25	30	
Ala Ala Gly Val Val Ile Ala Ser Pro Ser Arg Thr Asp Pro Pro Tyr			
35	40	45	
Phe Phe Thr Trp Thr Arg Asp Ala Ala Leu Val Leu Thr Gly Ile Ile			
50	55	60	
Glu Ser Leu Gly His Asn Tyr Asn Thr Thr Leu Gln Thr Val Ile Gln			
65	70	75	80
Asn Tyr Val Ala Ser Gln Ala Lys Leu Gln Gln Val Ser Asn Pro Ser			

85	90	95
Gly Thr Phe Ala Asp Gly Ser Gly	Leu Gly Glu Ala Lys	Phe Asn Val
100	105	110
Asp Leu Thr Ala Phe Thr Gly Glu Trp Gly Arg Pro Gln Arg Asp Gly		
115	120	125
Pro Pro Leu Arg Ala Ile Ala Leu Ile Gln Tyr Ala Lys Trp Leu Ile		
130	135	140
Ala Asn Gly Tyr Lys Ser Thr Ala Lys Ser Val Val Trp Pro Val Val		
145	150	155
Lys Asn Asp Leu Ala Tyr Thr Ala Gln Tyr Trp Asn Glu Thr Gly Phe		
165	170	175
Asp Leu Trp Glu Glu Val Pro Gly Ser Ser Phe Phe Thr Ile Ala Ser		
180	185	190
Ser His Arg Ala Leu Thr Glu Gly Ala Tyr Leu Ala Ala Gln Leu Asp		
195	200	205
Thr Glu Cys Arg Ala Cys Thr Thr Val Ala Pro Gln Val Leu Cys Phe		
210	215	220
Gln Gln Ala Phe Trp Asn Ser Lys Gly Asn Tyr Val Val Ser Asn Ile		
225	230	235
Asn Gly Gly Glu Tyr Arg Ser Gly Lys Asp Ala Asn Ser Ile Leu Ala		
245	250	255
Ser Ile His Asn Phe Asp Pro Glu Ala Gly Cys Asp Asn Leu Thr Phe		
260	265	270
Gln Pro Cys Ser Glu Arg Ala Leu Ala Asn His Lys Ala Tyr Val Asp		
275	280	285
Ser Phe Arg Asn Leu Tyr Ala Ile Asn Lys Gly Ile Ala Gln Gly Lys		
290	295	300
Ala Val Ala Val Gly Arg Tyr Ser Glu Asp Val Tyr Tyr Asn Gly Asn		
305	310	315
Pro Trp Tyr Leu Ala Asn Phe Ala Ala Ala Glu Gln Leu Tyr Asp Ala		
325	330	335
Ile Tyr Val Trp Asn Lys Gln Gly Ser Ile Thr Val Thr Ser Val Ser		
340	345	350
Leu Pro Phe Phe Arg Asp Leu Val Ser Ser Val Ser Thr Gly Thr Tyr		
355	360	365
Ser Lys Ser Ser Ser Thr Phe Thr Asn Ile Val Asn Ala Val Lys Ala		
370	375	380
Tyr Ala Asp Gly Phe Ile Glu Val Ala Ala Lys Tyr Thr Pro Ser Asn		
385	390	395
Gly Ala Leu Ala Glu Gln Tyr Asp Arg Asn Thr Gly Lys Pro Asp Ser		
405	410	415
Ala Ala Asp Leu Thr Trp Ser Tyr Ser Ala Phe Leu Ser Ala Ile Asp		
420	425	430
Arg Arg Ala Gly Leu Val Pro Pro Ser Trp Arg Ala Ser Val Ala Lys		
435	440	445
Ser Gln Leu Pro Ser Thr Cys Ser Arg Ile Glu Val Ala Gly Thr Tyr		
450	455	460
Val Ala Ala Thr Ser Thr Ser Phe Pro Ser Lys Gln Thr Pro Asn Pro		
465	470	475
Ser Ala Ala Pro Ser Pro Ser Pro Tyr Pro Thr Ala Cys Ala Asp Ala		
485	490	495
Ser Glu Val Tyr Val Thr Phe Asn Glu Arg Val Ser Thr Ala Trp Gly		
500	505	510
Glu Thr Ile Lys Val Val Gly Asn Val Pro Ala Leu Gly Asn Trp Asp		
515	520	525
Thr Ser Lys Ala Val Thr Leu Ser Ala Ser Gly Tyr Lys Ser Asn Asp		
530	535	540

Pro	Leu	Trp	Ser	Ile	Thr	Val	Pro	Ile	Lys	Ala	Thr	Gly	Ser	Ala	Val
545				550					555						560
Gln	Tyr	Lys	Tyr	Ile	Lys	Val	Gly	Thr	Asn	Gly	Lys	Ile	Thr	Trp	Glu
				565				570						575	
Ser	Asp	Pro	Asn	Arg	Ser	Ile	Thr	Leu	Gln	Thr	Ala	Ser	Ser	Ala	Gly
					580			585							590
Lys	Cys	Ala	Ala	Gln	Thr	Val	Asn	Asp	Ser	Trp	Arg				
				595			600								

<210> 4
<211> 10739
<212> DNA
<213> Artificial Sequence

<220>
<223> pTrex3g_N13 plasmid

<400> 4	aagcttacta	gtacttctcg	agctctgtac	atgtccggtc	gcgacgtacg	cgtatcgatg	60
	gcgccagctg	caggcggccg	cctgcagcca	cttgcagtcc	cgtggattc	tcacggtgaa	120
	tgttaggcctt	ttgttagggta	ggaatttgtca	ctcaagcacc	cccaacctcc	attacgcctc	180
	ccccatagag	ttcccaatca	gtgagtcatg	gcactgttct	caaatacgatt	ggggagaagt	240
	tgacttccgc	ccagagctga	aggtcgcaca	accgcatgtat	atagggtcgg	caacggcaaa	300
	aaagcacgtg	gctcaccgaa	aagcaagatg	tttgcgatct	aacatccagg	aacctggata	360
	catccatcat	cacgcacgac	cactttgatc	tgctggtaaa	ctcgtattcg	ccctaaaccg	420
	aagtgcgtgg	taaatctaca	cgtgggcccc	tttcggata	ctgcgtgtgt	cttctctagg	480
	tgccatttctt	ttcccttccct	ctagtgttga	attgtttgtg	ttggagtcgg	agctgttaact	540
	acctctgaat	ctctggagaa	tggtggacta	acgactaccg	tgcacctgca	tcatgtatat	600
	aatagtgatc	ctgagaaggg	gggtttggag	caatgtggga	ctttgatggt	catcaaacaa	660
	agaacgaaga	cgcctctttt	gcaaagtttt	gttccggcta	cggtgaagaa	ctggataactt	720
	gttgcgtctt	ctgtgtatTTT	ttgtggcaac	aagaggccag	agacaatcta	ttcaaaccacc	780
	aagcttgctc	ttttgagcta	caagaacctg	tggggtatata	atctagagtt	gtgaagtcgg	840
	taatcccgt	gtatagtaat	acgagtcgca	tctaaatact	ccgaagctgc	tgcgaacccg	900
	gagaatcggag	atgtgctgga	aagcttctag	cgagccgcta	aattagcatg	aaaggctatg	960
	agaaaattctg	gagacggctt	gttgaatcat	ggcggtccat	tcttcgacaaa	gcaaagcggt	1020
	ccgtcgcagt	agcaggcact	cattcccgaa	aaaactcgga	gattcctaag	tagcgatgg	1080
	accggaataaa	tataataggc	aatacattga	gttgcctcga	cggttgcatt	gcagggggtac	1140
	ttagcttggaa	cataactgtt	ccgtacccca	cctcttctca	acctttggcg	tttccctgat	1200
	tcagcgtacc	cgtacaagtgc	gtaatcacta	ttaacccaga	ctgaccggac	gtgtttgcc	1260
	cttcatttgg	agaaataatg	tcattgcgtat	gtgttaatttgc	cctgcttgcac	cgactggggc	1320
	tgttcgaagc	ccgaatgttag	gattgttatac	cgaactctgc	tcgttagagggc	atgttgcgaa	1380
	tctgtgtcgg	gcaggacacg	cctcgaagggt	tcacggcaag	ggaaaccacc	gatagcagt	1440
	tctagtagca	acctgttaaag	ccgcaatgca	gcatcactgg	aaaatacaaa	ccaatggcta	1500
	aaagtacata	agttaatgcc	taaagaagtc	atataccaggc	ggctaataat	tgtacaatca	1560
	atggctaaa	cgtaccgtaa	tttgccaaacg	gcttgtgggg	ttgcagaagc	aacggcaaag	1620
	ccccacttcc	ccacgtttgt	tttccactc	agtccaaatct	cagctggtga	tcccccaatt	1680
	gggtcgcttgc	tttgcgtccgg	tgaagtggaaa	gaagacagag	gtaagaatgt	ctgactcgga	1740
	gcgttttgca	tacaaccaag	ggcagtgtat	gaagacagtg	aaatgttgcac	attcaaggag	1800
	tatttagcca	gggatgtttgc	agtgtatcgat	gtaaggaggt	ttgtctggcc	atacgacgaa	1860
	tactgtatag	tcacttctga	tgaagtggtc	catattgaaa	tgtaaagtcg	gcactgaaca	1920
	ggcaaaagat	tgagttggaaa	ctgcctaaga	tctcgggcccc	tcggggccttc	ggcctttggg	1980
	tgtacatgtt	tgtgcgtccgg	gcaaatacgaa	agtgtggtag	gatcgaaacac	actgctgcct	2040
	ttaccaagca	gctgagggttgc	tgtgataggc	aaatgttgcag	ggccactgc	atgggttcga	2100
	atagaaagag	aagcttagcc	aagaacaata	gccgataaag	atagcctcat	taaacggaaat	2160
	gagcttagtag	gcaaagtcag	cgaatgtgt	tatataaagg	ttcgaggtcc	gtgcctccct	2220
	catgctctcc	ccatctactc	atcaactcag	atcctccagg	agacttgcac	accatctttt	2280
	gaggcacaga	aaccaatag	tcaaccatca	caagttgt	caaaaaagca	ggctccgcgg	2340

ccgccccctt	caacatgcat	accttctcca	agctcctcg	cctgggctct	gccgtccagt	2400
ctgcctcgg	gccccctcac	ggcttctcg	gtctccagga	acgcgcgtgc	gttgatacct	2460
tcatcaacac	cgagaagccc	atcgcatgga	acaagctgct	cgccaaacatc	ggccctaacg	2520
gcaaagccgc	tcccggtgcc	gccgccccgg	ttgtgattgc	cagcccttcc	aggacggacc	2580
ctccttgtac	gtggtggcat	ggaatggacc	caagagactg	gttttagatg	aaagagagtt	2640
tctgctaacc	gccacaccca	gacttctca	cctggacc	cgatgcgc	ctggcctca	2700
ccggcatcat	cgagtcctt	ggccacaact	acaacaccac	cctgcagacc	gtcatccaga	2760
actacgtcg	gtgcgcaggcc	aagctgcagc	aggtctcgaa	cccctcgaaa	accttcgc	2820
acggctcg	tctcggtgag	gccaagttca	atgtcgac	cactgccttc	actggcgaat	2880
ggggtcgccc	tcagagggac	ggcccgcccc	tgcgcgc	cgctctc	cagtacgcca	2940
agtggctgat	cgccaaacggc	tacaagagca	cgccaaagag	cgtcgtctgg	cccgtcgta	3000
agaacgatct	cgcctacacg	gcccagtact	ggaacgagac	cggttcgat	ctctgggagg	3060
aggtccccgg	cagctcg	tttaccatcg	ccagctctca	cagggtgtgag	tcatttattg	3120
ttcagtgtt	tctcattgaa	taattaccgg	aatgcactg	acgcca	gctctgactg	3180
agggtgctt	cctcgccg	cagctcgaca	ccgagtgcc	cgccctgcac	accgtcgccc	3240
ctcaggttct	gtgctccag	caggccttct	ggaactccaa	gggcaactat	gtcgctcca	3300
acagtaagat	ccctacacca	acaaaaaaa	tcgaaaagga	acgttagctg	acccttctag	3360
tcaacggcgg	cgagtatcg	tccggcaagg	acgccaactc	gatcctgg	tccatccaca	3420
acttcgaccc	tgaggccgg	tgcgacaacc	tgacctcca	gccctgcagc	gagcgcgccc	3480
tggccaacca	caaggcctat	gtcgactcg	tccgcaac	ctacgccc	aacaaggca	3540
tcgcccagg	caaggccgtt	gccgtcg	gctactcg	ggatgtctac	tacaacggca	3600
accctggta	cctggccaac	tttgcgc	ccgagcag	ctacgac	atctacgtgt	3660
ggaacaagca	gggctccatc	accgtgac	cggtctcc	gcccttctt	cgcgac	3720
tctcgctg	cagcacccgg	actactca	agagcag	gac	tttaccatc	3780
acggcgtaa	ggcctacg	gacggcttca	tcgagg	ggcca	acccgttca	3840
acggcg	cgccgagc	tacgaccgca	acacgg	gccc	actcg	3900
tgacgtgg	gtactcg	ttcctctcg	ccatcgac	ccgcgcgg	ctcg	3960
cgagctgg	ggccagcgt	gccaagag	agctcg	cac	ctcg	4020
tcgcccgc	ctacgtcg	gccacg	cctcg	gtcc	aa	4080
cctccgcgc	gccctcc	tccc	tacc	cgac	cg	4140
acgtcac	caacgagc	gtgtcgac	cgtgg	gacc	atcg	4200
acgtccgg	gctggg	tgac	ccag	ctgt	gtgg	4260
acaagtgc	tgatccc	tggacat	cggt	caagg	ggcc	4320
tgcagtaca	gtat	atca	gtcg	gacca	tcttgg	4380
acaggagcat	tacc	ctcg	acgg	gtcg	ggcc	4440
atgattcg	gcgtt	aaa	gtgg	ccg	cg	4500
tcgcgc	tccgt	gcgaa	agc	ttt	ttt	4560
gacggcgg	ggag	ctacat	ggcc	ttt	ttt	4620
ctttcaat	atacgg	ctcat	actgg	ttt	ttt	4680
tgttgc	ttgg	caaatt	gtgg	ttt	ttt	4740
catttaaga	taacgg	ata	gaag	aaa	aaa	4800
ccgttcaaa	ccgt	tagaa	ttt	ttt	ttt	4860
agctcg	ctgg	agag	tc	ttt	ttt	4920
gttgtc	ggag	cg	ttt	ttt	ttt	4980
atgttgc	gcagg	ctg	ttt	ttt	ttt	5040
aatcgc	ggga	aggcc	ttt	ttt	ttt	5100
atcaga	cacgt	aaactcg	ttt	ttt	ttt	5160
ccagtgc	gtt	tctgc	ttt	ttt	ttt	5220
ccagctagg	acc	gtt	ttt	ttt	ttt	5280
ggatca	gg	gaaat	ttt	ttt	ttt	5340
aactcc	tt	acactcg	ttt	ttt	ttt	5400
aatcc	aga	actgg	ttt	ttt	ttt	5460
aatgg	cc	agacg	ttt	ttt	ttt	5520
ggatc	agagg	ccg	ttt	ttt	ttt	5580
agctgg	cg	gag	ttt	ttt	ttt	5640
caatcg	gc	agtt	ttt	ttt	ttt	5700
ttatgg	atca	agctg	ttt	ttt	ttt	5760

ccctgacgcc	gctctcgccg	aggcaaggga	actcgatgaa	tactacgcaa	agcacaagag	5820
accgggtgt	ccactccatg	gcctccccat	ctctctcaa	gaccagctt	gagtcaaggt	5880
acaccgtgc	ccctaagtcg	ttagatgtcc	cttttgtca	gctaaacatat	gccaccaggg	5940
ctacgaaaca	tcaatgggct	acatctcatg	gctaaacaag	tacgacgaag	gggactcggt	6000
tctgacaacc	atgctccgca	aagccggtgc	cgtcttctac	gtcaagacct	ctgtcccgca	6060
gaccctgatg	gtctgcgaga	cagtcaacaa	catcatcggg	cgcaccgtca	acccacgcaa	6120
caagaactgg	tcgtgcggcg	gcagttctgg	tggtgagggt	gcatcggtt	ggattcgcrv	6180
tggtggcgctc	atcggtgttag	gaacggatat	cggtggctcg	attcgagtgc	cgcccgctt	6240
caacttcctg	tacggctcaa	ggccgagtca	tgggcccgtg	ccgtatgcaa	agatggcgaa	6300
cagcatggag	ggtcaggaga	cgggtcacak	cgttgcggg	ccgattacgc	actctgttga	6360
gggtgagtcc	ttcgccctt	ccttctttc	ctgctctata	ccaggccctcc	actgtcctcc	6420
tttcttgctt	tttatactat	atacgagacc	ggcagtcact	gatgaagtat	gttagacctc	6480
cgccttcata	ccaaatccgt	cctcggtcag	gagccatgga	aatacgactc	caaggtcata	6540
cccatgcct	ggcgcctcagtc	cgagtccggac	attattgcct	ccaagatcaa	gaacggcggg	6600
ctcaatatcg	gctactacaa	cttcgacggc	aatgtccttc	cacaccctcc	tatcctgcgc	6660
ggcgtggaaa	ccaccgtcgc	cgactcgcc	aaagccggc	acaccgtgac	cccggtgacg	6720
ccatacaagc	acgatttcgg	ccacgatctc	atctccata	tctacgcggc	tgacggcagc	6780
crvgccgacg	taatgcgcga	tatcagtgc	tccggcgagc	ccgcgattcc	aaatatcaa	6840
gacctactga	acccgaacat	caaagctgtt	aacatgaacg	agctctggga	cacgatctc	6900
cagaagtgg	attaccagat	ggagtacctt	gagaaatggc	gggaggctg	agaaaaaggcc	6960
gggaaggaac	tggacccat	catcgcccg	attacgccta	ccgctcggt	acggcatgac	7020
cagttccgg	actatgggt	tgccctgtt	atcaacctgc	tggttgcac	gagcgtgggt	7080
gttccggta	ccttgcgga	taagaacatc	gataagaaga	atgagagtt	caaggccgtt	7140
agttagctt	atgcctcgt	gcaggaagag	tatgatccgg	aggcgtacca	tggggcaccg	7200
gttgcagtgc	agtttatcgg	acggagactc	agtgaagaga	ggacgttggc	gattcagag	7260
gaagtgggaa	agttgctgg	aatatgtgt	actccatagc	taataagtgt	cagatagcaa	7320
tttgcacaag	aaatcaatac	cagcaactgt	aaataagcgc	tgaagtgacc	atgcctatgt	7380
acgaaaagagc	agaaaaaaaaac	ctgcccgtaga	accgaagaga	tatgacacgc	ttccatctct	7440
caaaggaaga	atcccttcag	ggttgcgtt	ccagtcata	cacgtataac	ggcacaagtg	7500
tctctcacca	aatgggttat	atctcaatg	tgtatctt	atggaaagcc	cagaatatcg	7560
atcgccgcga	gatccatata	tagggcccg	gttataatta	cctcaggtcg	acgtcccatg	7620
gccattcgaa	ttcgtatca	tggtcata	tgtttctgt	gtgaaattgt	tatccgtca	7680
caattccaca	caacatacga	gccggaaagca	taaagtgtaa	agcctgggg	gcctaatgag	7740
tgagctaact	cacattaatt	gcgttgcgt	cactgcccgc	tttccagtcg	ggaaaacctgt	7800
cgtgccagct	gcattaatga	atcggccaaac	gcccggggag	aggcgttgc	cgtattggc	7860
gctttccgc	ttcctcgctc	actgactcgc	tgcgtcggt	cggtcggt	cgccgagcgg	7920
tatcagctca	ctcaaaggcg	gtaatacgg	tatccacaga	atcaggggat	aacgcaggaa	7980
agaacatgtg	agcaaaaaggc	cagcaaaagg	ccaggaaccg	taaaaaggcc	gcgttgcgg	8040
cgttttcca	taggctccgc	ccccctgacg	agcatcaca	aaatcgacgc	tcaagtca	8100
ggtggcgaaa	cccgacacg	ctataaagat	accaggcg	tcccccgtt	agctccctcg	8160
tgcgtctcc	tgttccgacc	ctgcccgtt	ccggatacct	gtccgcctt	ctcccttcgg	8220
gaagcgtggc	gtttctcat	agctcacgt	gttagtatct	cagttcggt	taggtcggt	8280
gctccaagct	gggctgtgt	cacgaacccc	ccgttcagcc	cgaccgctgc	gccttatccg	8340
gtaactatcg	tcttgagtcc	aacccggtaa	gacacgactt	atcgccactg	gcagcagcc	8400
ctggtaacag	gattagcaga	gcgaggtatg	taggcggtgc	tacagagtgc	ttgaagtgg	8460
ggcttaacta	cggctacact	agaagaacag	tatttggat	ctgcgtctg	ctgaagccag	8520
ttaccttcgg	aaaaagagtt	ggtagctt	gatccggcaa	acaaaccacc	gctggtagcg	8580
gtggttttt	tgttgcag	cagcagatta	cgccgagaaa	aaaaggatct	caagaagatc	8640
ctttgatctt	ttctacgggg	tctgacgtc	agtggaaacg	aaactcacgt	taagggattt	8700
tggtcatgag	attatcaaaa	aggatctca	cctagatctt	tttaaattaa	aatagaagtt	8760
ttaaatcaat	ctaaagtata	tatgagtaa	cttggcttgc	cagttacca	tgcttaatca	8820
gtgaggcacc	tatctcagcg	atctgtctat	ttcgttcatc	catagttgc	tgactccccg	8880
tcgtgttagat	aactacgata	cgggaggggct	taccatctgg	ccccagtgc	gcaatgatac	8940
cgcgagaccc	acgctcaccg	gctccagatt	tatcagcaat	aaaccagcca	gccggaaagg	9000
ccgagcgcag	aagtggctt	gcaactttat	ccgcctccat	ccagtctatt	aatttttgcc	9060
gggaagctag	agtaagtagt	tcgcccgtt	atagttgc	caacgttgc	gccattgcta	9120
caggcatcg	ggtgtcacgc	tcgtcggtt	gtatggctt	attcagctcc	ggttcccaac	9180

gatcaaggcg	agttacatga	tccccatgt	tgtgaaaaaa	agcggttagc	tcctcggtc	9240
ctccgatgt	tgtcagaagt	aagtggccg	cagtgttata	actcatggtt	atggcagcac	9300
tgcataattc	tcttactgtc	atgccatccg	taagatgctt	ttctgtgact	ggtgagttact	9360
caaccaagtc	attctgagaa	tagtgtatgc	ggcgaccgag	ttgcttgc	ccggcgtcaa	9420
tacgggataa	taccgcgcca	catagcagaa	cttaaaagt	gctcatcatt	ggaaaacgtt	9480
cttcggggcg	aaaactctca	aggatcttac	cgctgtttag	atccagttcg	atgttaaccca	9540
ctcggtcacc	caactgtatct	tcagcatctt	ttactttcac	cagcgtttct	gggtgagcaa	9600
aaacaggaag	gaaaaatgcc	gcaaaaaaagg	gaataaggc	gacacggaaa	tgttgaatac	9660
tcatactctt	ccttttcaa	tattattgaa	gcatttatca	gggttattgt	ctcatgagcg	9720
gatacatatt	tgaatgtatt	tagaaaaata	aacaaatagg	ggttccgcgc	acatttcccc	9780
gaaaagtgcc	acctgacgtc	taagaaacca	ttattatcat	gacattaacc	tataaaaata	9840
ggcggtatcac	gaggccctt	cgtctcgcc	gttgcgttga	tgacggtga	aacctctgac	9900
acatgcagct	cccggagacg	gtcacagctt	gtctgttgc	ggatgcccgg	agcagacaag	9960
cccggtcagg	cgcgtcagcg	ggtgttgcgc	ggtgtccggg	ctggcttaac	tatgcggcat	10020
cagagcagat	tgtactgaga	gtgcaccata	aaattgtaaa	cgttaatatt	ttgttaaaat	10080
tcgcgttaaa	tttttgttaa	atcagctcat	tttttaacca	ataggccgaa	atcgccaaaa	10140
tcccttataa	atcaaaagaa	tagcccgaga	taggtttag	tgttgttcca	gttttggaaaca	10200
agagtccact	attaaagaac	gtggactcca	acgtcaaagg	gcaaaaaacc	gtctatcagg	10260
gcatggccc	actacgtgaa	ccatcaccca	aatcaagttt	tttggggcgt	aggtgccgt	10320
aagactaaa	tccggAACCT	aaagggagcc	cccgatttag	agcttgcacgg	ggaaaagccgg	10380
cgaacgtggc	gagaaaggaa	gggaagaaag	cgaaaggagc	gggcgtttag	gctggcaaa	10440
gtgttagcggt	cacgcgtcgc	gttaaccacca	cacccgcgc	gcttaatgcg	ccgctacagg	10500
gcccgtacta	tgggtgttt	gacgtatgc	gtgtgaaata	ccgcacagat	gcttaaggag	10560
aaaataccgc	atcaggcgcc	attcgccatt	caggctcgcc	aactgttggg	aaggcgatc	10620
gtgtcgccgc	tcttcgttat	tacgcccagct	ggcgaaagg	ggatgtgct	caaggcgatt	10680
aagttggta	acgcccagg	tttcccagtc	acgacgttgt	aaaacgacgg	ccagtgccc	10739

<210> 5
<211> 2169
<212> DNA
<213> Aspergillus awamori var. kawachi

<400> 5						
atgtcggtcc	gatcttttct	cgccctgagc	ggccttgc	gctcggtt	ggcaagtgt	60
atttccaagc	gcgcgacctt	ggattcgtgg	ttgagcaacg	aagcgaccgt	ggcccgact	120
gcgcgttgc	ataacatcg	ggcgacgg	gcttgggtgt	cgccgcgc	ctctggcatt	180
gtcggttgc	gtcccagcac	cgataacccg	gactgtatgt	tttgagttcg	gattatgaat	240
gtgtcttgc	tgattgtatgc	tgactggcg	gtcttttgc	gattgttagac	ttctacacct	300
ggactcgcga	ctctggtctc	gtcatcaaga	ccctcgctc	ccttccgc	aatggagata	360
ctgatctcct	ttccaccatt	gagcaactaca	tctccctca	ggcaattatt	cagggtgtca	420
gtaaaccctc	tggtgatctg	tccagcggt	gtcttgc	gccaagttc	aatgtcgatg	480
agactgccta	caccggttct	tggggacggc	cgcagctg	tggctctgc	ctgagagcaa	540
ctgctatgt	cggttttgg	cagtggtgc	ttgtatgtt	tccacctcct	tgcgtctgat	600
ctgcaacata	tgtagccgac	tggtcaggac	aatggctaca	ccagcgctgc	aacagagatt	660
gtttggccccc	tgcgtttagaa	cgacctgtcg	tatgtggctc	agtaactggaa	ccagacggga	720
tatgggtgt	ttgattgtatc	gggggttcaag	ggtgtttgt	catcgagct	aacttcgcgg	780
tcgcagatct	ctgggaagaa	gttaatggct	cgcccttctt	cactattgcc	gtcaacacc	840
gcgcgcgt	cgaaggtagt	gccttcgc	cgccgcgt	ctcgctctc	tcctgggt	900
attcgcaggc	acctcagatt	ctctgttact	tgcgttctt	ctggaccggc	agctacatcc	960
tggccaaactt	tgacagcagc	cggtccggc	aggacacaaa	caccctcct	ggaagcatcc	1020
acaccttgc	tcctgaggct	ggatgcgacg	actccaccc	ccagccctgc	tcccccgcgt	1080
cgctcgccaa	ccataaggag	gtttagact	cttccgc	gatctatact	ctcaacgatg	1140
gtctcgttgc	cagtggcg	gttgcgg	gtcggtaccc	tgaggatagc	tactacaacg	1200
gcaacccgt	gttcctgtc	accttggct	ccgcggaaaca	gctgtacgt	gctctgtacc	1260
agtgggacaa	gcaggggtcg	ttggagatca	cagacgtgc	acttgactt	ttcaaggctc	1320
tgtacagtgg	tgctgcccacc	ggcacgtact	cttcgtccag	ctcgacctat	agcagcattg	1380
tgagtggccgt	caagactt	gctgtatgg	ttgtttctat	tgtgttaagt	ctacgctaga	1440

cgagcgctca	tat	ttacaga	gggtgcgtac	taacaggatt	aggaaactca	cggcga	1500
aacggcttc	tgt	ctgagca	attcgacaag	tctgatggcg	acgagcttc	tgctcgat	1560
ctgacctgtt	ctt	acgctgc	tctgctgacc	gccaacaacc	gtcgtaattc	tgtcg	1620
ccgtcttggg	gtg	agacctc	tgccagc	gtgcccggca	cctgtgcggc	tacctctg	1680
tctggta	acct	acagcagt	gaccgtc	acc	tcgtggccga	gcatcg	1740
accactacg	cg	gctactac	cactggatcg	ggcggcgtga	cctcgaccag	caagaccacc	1800
acaactgct	taa	gaccag	caccactacg	tcctcgac	cctgcaccac	ccccactg	1860
gtagctgt	cctt	gatct	gacggcgacc	accac	ctcgatct	ctac	1920
gggtcgatct	ctc	agctcg	tgactggag	acc	cgatcgat	gagcg	1980
aagtacactt	ccag	caaccc	gcttggat	gtaactgt	ctctg	ccggc	2040
tttgagtaca	at	ttcatccg	cgtcgag	gatgactccg	tgagtg	gagcgacccg	2100
aaccggaa	at	accgttcc	tcaggcgt	ggcgagtc	ccgcgac	gaccgacacc	2160
tggcgtag							2169

<210> 6
<211> 24
<212> PRT
<213> Aspergillus awamori var. kawachi

<400>	6															
Met	Ser	Phe	Arg	Ser	Leu	Leu	Ala	Leu	Ser	Gly	Leu	Val	Cys	Ser	Gly	
1																15
Leu	Ala	Ser	Val	Ile	Ser	Lys	Arg									
																20

<210> 7
<211> 615
<212> PRT
<213> Aspergillus awamori var. kawachi

<400>	7															
Ala	Thr	Leu	Asp	Ser	Trp	Leu	Ser	Asn	Glu	Ala	Thr	Val	Ala	Arg	Thr	
1																15
Ala	Ile	Leu	Asn	Asn	Ile	Gly	Ala	Asp	Gly	Ala	Trp	Val	Ser	Gly	Ala	
																20
Asp	Ser	Gly	Ile	Val	Val	Ala	Ser	Pro	Ser	Thr	Asp	Asn	Pro	Asp	Tyr	
																35
Phe	Tyr	Thr	Trp	Thr	Arg	Asp	Ser	Gly	Leu	Val	Ile	Lys	Thr	Leu	Val	
																50
Asp	Leu	Phe	Arg	Asn	Gly	Asp	Thr	Asp	Leu	Leu	Ser	Thr	Ile	Glu	His	
																65
Tyr	Ile	Ser	Ser	Gln	Ala	Ile	Ile	Gln	Gly	Val	Ser	Asn	Pro	Ser	Gly	
																85
Asp	Leu	Ser	Ser	Gly	Gly	Leu	Gly	Glu	Pro	Lys	Phe	Asn	Val	Asp	Glu	
																100
Thr	Ala	Tyr	Thr	Gly	Ser	Trp	Gly	Arg	Pro	Gln	Arg	Asp	Gly	Pro	Ala	
																115
Leu	Arg	Ala	Thr	Ala	Met	Ile	Gly	Phe	Gly	Gln	Trp	Leu	Leu	Asp	Asn	
																130
Gly	Tyr	Thr	Ser	Ala	Ala	Thr	Glu	Ile	Val	Trp	Pro	Leu	Val	Arg	Asn	
																145
Asp	Leu	Ser	Tyr	Val	Ala	Gln	Tyr	Trp	Asn	Gln	Thr	Gly	Tyr	Asp	Leu	
																165
Trp	Glu	Glu	Val	Asn	Gly	Ser	Ser	Phe	Phe	Thr	Ile	Ala	Val	Gln	His	
																180
Arg	Ala	Leu	Val	Glu	Gly	Ser	Ala	Phe	Ala	Thr	Ala	Val	Gly	Ser	Ser	
																195
																200
																205

Cys Ser Trp Cys Asp Ser Gln Ala Pro Gln Ile Leu Cys Tyr Leu Gln
 210 215 220
 Ser Phe Trp Thr Gly Ser Tyr Ile Leu Ala Asn Phe Asp Ser Ser Arg
 225 230 235 240
 Ser Gly Lys Asp Thr Asn Thr Leu Leu Gly Ser Ile His Thr Phe Asp
 245 250 255
 Pro Glu Ala Gly Cys Asp Asp Ser Thr Phe Gln Pro Cys Ser Pro Arg
 260 265 270
 Ala Leu Ala Asn His Lys Glu Val Val Asp Ser Phe Arg Ser Ile Tyr
 275 280 285
 Thr Leu Asn Asp Gly Leu Ser Asp Ser Glu Ala Val Ala Val Gly Arg
 290 295 300
 Tyr Pro Glu Asp Ser Tyr Tyr Asn Gly Asn Pro Trp Phe Leu Cys Thr
 305 310 315 320
 Leu Ala Ala Ala Glu Gln Leu Tyr Asp Ala Leu Tyr Gln Trp Asp Lys
 325 330 335
 Gln Gly Ser Leu Glu Ile Thr Asp Val Ser Leu Asp Phe Phe Lys Ala
 340 345 350
 Leu Tyr Ser Gly Ala Ala Thr Gly Thr Tyr Ser Ser Ser Ser Ser Thr
 355 360 365
 Tyr Ser Ser Ile Val Ser Ala Val Lys Thr Phe Ala Asp Gly Phe Val
 370 375 380
 Ser Ile Val Glu Thr His Ala Ala Ser Asn Gly Ser Leu Ser Glu Gln
 385 390 395 400
 Phe Asp Lys Ser Asp Gly Asp Glu Leu Ser Ala Arg Asp Leu Thr Trp
 405 410 415
 Ser Tyr Ala Ala Leu Leu Thr Ala Asn Asn Arg Arg Asn Ser Val Val
 420 425 430
 Pro Pro Ser Trp Gly Glu Thr Ser Ala Ser Ser Val Pro Gly Thr Cys
 435 440 445
 Ala Ala Thr Ser Ala Ser Gly Thr Tyr Ser Ser Val Thr Val Thr Ser
 450 455 460
 Trp Pro Ser Ile Val Ala Thr Gly Gly Thr Thr Thr Ala Thr Thr
 465 470 475 480
 Thr Gly Ser Gly Gly Val Thr Ser Thr Ser Lys Thr Thr Thr Ala
 485 490 495
 Ser Lys Thr Ser Thr Thr Ser Ser Thr Ser Cys Thr Thr Pro Thr
 500 505 510
 Ala Val Ala Val Thr Phe Asp Leu Thr Ala Thr Thr Tyr Gly Glu
 515 520 525
 Asn Ile Tyr Leu Val Gly Ser Ile Ser Gln Leu Gly Asp Trp Glu Thr
 530 535 540
 Ser Asp Gly Ile Ala Leu Ser Ala Asp Lys Tyr Thr Ser Ser Asn Pro
 545 550 555 560
 Leu Trp Tyr Val Thr Val Thr Leu Pro Ala Gly Glu Ser Phe Glu Tyr
 565 570 575
 Lys Phe Ile Arg Val Glu Ser Asp Asp Ser Val Glu Trp Glu Ser Asp
 580 585 590
 Pro Asn Arg Glu Tyr Thr Val Pro Gln Ala Cys Gly Glu Ser Thr Ala
 595 600 605
 Thr Val Thr Asp Thr Trp Arg
 610 615

<210> 8
 <211> 28
 <212> DNA
 <213> Artificial Sequence

```
<220>
<223> primer

<400> 8
caacatgcat actttctcca agtcctc 28

<210> 9
<211> 25
<212> DNA
<213> Artificial Sequence

<220>
<223> primer

<400> 9
ttaacgccac gaatcattca ccgtc 25

<210> 10
<211> 25
<212> DNA
<213> Artificial Sequence

<220>
<223> primer

<400> 10
caccatgtcg ttccgatctc ttctc 25

<210> 11
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<223> primer

<400> 11
ctaccggccag gtgtcggtca c 21
```